Nonlinear Estimation with Perron-Frobenius Operator and Karhunen-Loeve Expansion

Parikshit Dutta, Abhishek Halder, Raktim Bhattacharya

▶ To cite this version:

HAL Id: hal-00773533

https://hal.inria.fr/hal-00773533

Submitted on 13 Aug 2014

Nonlinear Estimation with Perron-Frobenius Operator and Karhunen-Loève Expansion

Parikshit Dutta, *Member, IEEE,* Abhishek Halder, *Member, IEEE,*

and Raktim Bhattacharya, *Member, IEEE*

Abstract

In this paper, a novel methodology for state estimation of stochastic dynamical systems is proposed. In this formulation, finite-term Karhunen-Loéve (KL) expansion is used to approximate the process noise, thus resulting a non-autonomous deterministic approximation (with parametric uncertainty) of the original stochastic nonlinear system. It is proved that the solutions of the approximate dynamical system, asymptotically converge to the true solutions, in mean square sense. The evolution of uncertainty for the KL approximated system is predicted via Perron-Frobenius (PF) operator. Furthermore, a nonlinear estimation algorithm, using the proposed uncertainty propagation scheme, is developed. It is found that for finite dimensional linear and nonlinear filters, the evolving posterior densities obtained from the KLPF based estimator is closer, than particle filter, to the true posterior densities. The methodology is then applied to estimate states of a hypersonic reentry vehicle. It is observed that, the KLPF based estimator outperformed particle filter in terms of capturing localization of uncertainty through posterior densities, and reduction of uncertainty.

Index Terms

Karhunen-Loéve expansion, Perron-Frobenius operator, nonlinear estimation and filtering, hypersonic re-entry.

Parikshit Dutta is with the Department of Civil and Environmental Engineering, Duke University, Durham, NC 27708, USA, `parikshit.dutta@duke.edu`; Abhishek Halder is with Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA, `ahalder@tamu.edu`; and Raktim Bhattacharya is with the Department of Aerospace Engineering, Texas A&M University, College Station, TX 77843, USA, `raktim@tamu.edu`.

I. Introduction

Consider an uncertain nonlinear system given by the Itô stochastic differential equation (SDE)

$$dx(t) = f(x(t), t) \, dt + d\mathcal{W}(\omega, t), \tag{1}$$

$$dy(t) = h(x(t), t) \, dt + d\mathcal{V}(\omega, t), \tag{2}$$

where at time $t > 0$, the state vector $x(t) \in \mathbb{R}^n$, and the measurement vector $y(t) \in \mathbb{R}^m$. Further, $\mathcal{W}(\omega, t) : \Omega \times \mathbb{R}^+ \mapsto \mathbb{R}^n$, and $\mathcal{V}(\omega, t) : \Omega \times \mathbb{R}^+ \mapsto \mathbb{R}^m$ are mutually independent Wiener processes denoting process and measurement noise, respectively; where Ω denotes the sample space. The functions $f(.)$ and $h(.)$ represent the drift and measurement vector fields, respectively.

Solving the nonlinear estimation problem amounts to computing the belief or conditional probability density function (PDF) of the state vector $x(t)$, given the history of measurements till time t. Traditionally, this computation follows a two-step predictor-corrector structure. In the **first step**, a prediction of the state PDF is made using the dynamics model (1). If the uncertainty in initial condition ($x_0 := x(0)$) is prescribed through the joint PDF $\varphi_0(x_0)$, then the predicted state PDF $\varphi^-(x(t), t)$, also known as *prior PDF*, satisfies the *Fokker-Planck equation* (FPE) (or *Kolmogorov forward equation*) [1]

$$\frac{\partial \varphi^-}{\partial t} = -\sum_{i=1}^n \frac{\partial}{\partial x^{(i)}} \left(\varphi^- f_i \right) + \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 \left(Q^{(ij)} \varphi^- \right)}{\partial x^{(i)} \partial x^{(j)}} \tag{3}$$

subject to $\varphi^-(x(0), 0) = \varphi_0(x_0)$, where Q is the process noise covariance matrix. In the **second step**, the prior PDF $\varphi^-(x(t), t)$ obtained by solving (3), is updated using the measurement model (2) and Bayes' rule, to result the *posterior PDF* $\varphi^+(x(t), t)$.

A. Current State-of-the-art

Solving the parabolic partial differential equation (PDE) (3), second order in space and first order in time, is computationally hard [1]. Several methods [2] for approximate numerical solution of the FPE have been developed over the years. These algorithms, intend to solve the FPE using *either* grid based methods like FEM [3], *or* by using meshfree methods [4]. It is known [5] that for grid based methods, complexity in solving the problem increases exponentially with the increase in dimensions. Same is true [6], [7] for Monte Carlo based methods [8] for solving FPE. On the other hand, function approximation techniques like [9] have been reported to work well for moderate (3 to 4) dimensions. However, finding the correct basis to get a sufficiently good approximation of the transient PDFs remains challenging, particularly when the problem is high dimensional [10]. Thus, most solution methods for FPE suffer from the 'curse of dimensionality' [11]. As a result, nonlinear estimation schemes usually do not attempt to solve (3)

for computing the prior, instead works with the empirical estimate of the prior. This compromises the numerical accuracy of the overall estimation algorithm.

State estimation for nonlinear systems are commonly done using sequential Monte Carlo (SMC) methods, *particle filter* being the most popular amongst them [12]. Owing to the empirical estimate of the prior, these methods require a large number of ensembles for convergence [6], leading to high computational cost. This problem is tackled through *resampling* [13]. Particle filters with resampling technique, are commonly known as *bootstrap filters* [14]. It has been observed [15] that bootstrap filters introduce other problems like loss of diversity amongst particles, if the resampling is not performed correctly. Recently developed techniques [16] have combined importance sampling and Markov Chain Monte Carlo (MCMC) methods to improve sampling quality, enhancing the accuracy of the state estimates. However, even with resampling, due to the simulation based nature of these filters, the ensemble size scales exponentially with state dimension [17]. To circumvent this problem, particle filters based on Rao-Blackwellization [18] have been proposed to partially solve the estimation problem analytically. However, its application remains limited to systems where the required partition of the state space is possible.

B. Contribution of this Paper

In this work, we have proposed an uncertainty propagation methodology for computing the prior PDF, by combining Karhunen-Loève (KL) expansion [19] with Perron-Frobenius (PF) operator theory [20]. Further, we have developed a nonlinear state estimation technique which uses the proposed method for forward propagation of uncertainty. Notice that, in the absence of process noise, (3) reduces to *Liouville equation* associated with the PF operator [20] that describes spatio-temporal drift of $\varphi^-\left(x(t), t\right)$. Being a first order quasi-linear PDE, Liouville equation can be solved [21] using the method of characteristics (MOC) [22] and unlike Monte Carlo, such a computation occurs in pointwise exact arithmetic [7]. In nonlinear estimation setting, it has already been shown [23], [24] that for systems without process noise, PF operator based nonlinear filter outperforms particle filter and bootstrap filter. Hence, it is reasonable to hope that the proposed estimation technique based on parametric approximation of the process noise via KL expansion, and nonparametric uncertainty propagation via PF operator, will perform better than other SMC methods.

The motivation behind investigating such mixed parametric-nonparametric approach stems from the fact that the presence of process noise necessitates solving a second order transport PDE (3), and consequently, MOC based exact arithmetic computation can not be achieved. This argument usually leads researchers to develop approximation algorithms to solve the second order FPE, which is an exact description of the

problem. Here, instead, we derive an approximate ordinary differential equation (ODE) representation of the problem (using KL), and then solve this approximate problem in exact arithmetic (using PF). Hence our approach differs from those which strive to numerically solve the Fokker-Planck PDE, in the sense that we propose to approximate the problem while the latter strives to approximate the solution. Convincing numerical evidence in favor of this argument was given in our earlier work [25]. Through a multivariate Kolmogorov-Smirnov test, that paper also showed that the solution of the approximated problem satisfies *at least* weak distributional convergence to the true SDE solution, meaning our "first KL, then PF" (henceforth referred as KLPF) algorithm is asymptotically consistent in distribution. The formal proof of the consistency has appeared in [26]. However, it was not shown how the proposed KLPF filtering technique compares to Monte Carlo based techniques like particle filters in terms of minimizing the variance of estimation error and localization of uncertainty. This paper focuses on this issue of numerical performance of the estimation algorithm. However, two issues remained unsettled there: (i) in *exactly* what sense the convergence of the solution occurs, and (ii) formal proof in support of this consistency result. This paper addresses both these issues. Further, a detailed numerical investigation is carried out about the performance of the KLPF filter.

The main contribution of this paper is twofold: firstly, it has been shown numerically through application to the hypersonic entry problem that the proposed KLPF filter performs better than particle filters in terms of minimizing the estimation error variance. Secondly, we have also demonstrated that the KLPF filter captures the dispersion of uncertainty better than the Monte Carlo method based filters, like particle filters, when applied to the hypersonic entry problem.

C. Structure of this Paper

The rest of the paper is organized as follows. In **Section II**, KL expansion is introduced first, and used to approximate the process noise term. It is also shown why such an approximation yields a sufficiently good solution for the SDE. The KLPF methodology is presented next in **Section III** with some illustrative examples. Extension of this idea to other types of stochastic evolution equations is pointed out. Then the nonlinear estimation technique, based on KLPF, is proposed in **Section IV**. In **Section V**, we demonstrate that for finite dimensional linear and nonlinear filters with known solutions, KLPF filter outperforms particle filter. Then we apply this technique to estimate states of a six dimensional nonlinear system and compare its results with particle filter. **Section VI** concludes the paper.

II. KARHUNEN-LOÈVE APPROXIMATION OF STOCHASTIC DYNAMICAL SYSTEMS

KL expansion was derived independently by many researchers [19], [27] to represent a *second order centered stochastic process* $Y(\omega, t)$ as a *random* linear combination of a set of orthonormal deterministic $\mathcal{L}_2$ functions $\{e_i(t)\}_{i=1}^{\infty}$, i.e.

$$Y(\omega, t) = \sum_{i=1}^{\infty} Z_i(\omega) e_i(t), \tag{4}$$

where $Z_i(\omega)$ are *pairwise uncorrelated random variables*. The idea is similar to the Fourier series expansion, where a *deterministic* linear combination of orthonormal $\mathcal{L}_2$ functions is used.

Further, if we write $Z_i(\omega) = \sqrt{\Lambda_i} \zeta_i(\omega)$, where $\Lambda_i \in \mathbb{R}^+$, and $\{\zeta_i(\omega)\}_{i=1}^{\infty}$ is a sequence of random variables to be determined, then $\{\Lambda_i\}_{i=1}^{\infty}$ and $\{e_i(t)\}_{i=1}^{\infty}$ can be interpreted as the eigenvalues and eigenfunctions of the *covariance function*,

$$C(t_1, t_2) := \text{cov}(Y(\omega, t_1) - \mathbb{E}[Y(\omega, t_1)],$$

$$Y(\omega, t_2) - \mathbb{E}[Y(\omega, t_2)]) \quad, \tag{5}$$

that admits a spectral decomposition [28] of the form $C(t_1, t_2) = \sum_{i=1}^{\infty} \Lambda_i e_i(t_1) e_i(t_2)$. Since the covariance function is bounded, symmetric and positive-definite, the eigenvalue problem can be cast as a homogeneous Fredholm integral equation of second kind, given by,

$$\int_{\mathcal{D}_t} C(t_1, t_2) e_i(t_1) \, dt_1 = \lambda_i e_i(t_2). \tag{6}$$

Given the covariance function of a stochastic process, the eigenvalue-eigenfunction set can be found by solving the above equation, and the resulting expansion (4) converges to $Y(\omega, t)$ in mean-square sense [29]. The following results are useful for our formulation.

Theorem 1: (**KL expansion of Wiener process** [30]) For Wiener process $W(\omega, t)$, the eigenvalues and eigenfunctions $\{\lambda_i, e_i(t)\}_{i=1}^{\infty}$, of its covariance function $C(t_1, t_2) = \sigma^2(t_1 \wedge t_2)$, $t_1, t_2 \in [0, T] \times [0, T]$, are given by

$$\lambda_i = \frac{4T}{\pi^2 (2i-1)^2}, \qquad e_i(t) = \sqrt{\frac{2}{T}} \sigma \sin\left(\left(i - \frac{1}{2}\right) \frac{\pi t}{T}\right), \tag{7}$$

and hence the KL expansion of $W(\omega, t)$, is of the form

$$W(\omega, t) \stackrel{\text{m.s.}}{=} \sqrt{\frac{2}{T}} \sigma \sum_{i=1}^{\infty} \zeta_i(\omega) \frac{\sin\left(\left(i - \frac{1}{2}\right) \frac{\pi t}{T}\right)}{\left(i - \frac{1}{2}\right) \frac{\pi}{T}}, \tag{8}$$

where $\zeta_i(\omega)$ are i.i.d. samples drawn from $\mathcal{N}(0, 1)$.

Corollary 2: (**KL expansion of Gaussian white noise**) Since $dW(\omega, t) = \eta(\omega, t)\, dt$, the KL expansion for Gaussian white noise $\eta(\omega, t)$ can be obtained by taking the derivative of (8) with respect to t, i.e.

$$\eta(\omega, t) \overset{\text{m.s.}}{=} \sqrt{\frac{2}{T}} \sigma \sum_{i=1}^{\infty} \zeta_i(\omega) \cos\left(\left(i - \frac{1}{2}\right) \frac{\pi t}{T}\right), \tag{9}$$

where the i.i.d. random variables $\zeta_i(\omega) \sim \mathcal{N}(0, 1)$.

A. Langevin Form with Noise KL Expansion

We rewrite the stochastic dynamical system (1) in the Langevin form

$$\dot{x}(t) = f(x(t), t) + \eta(\omega, t). \tag{10}$$

We consider integrating up to a finite time T; hence, $\eta(\omega, t) : \Omega \times [0, T] \mapsto \mathbb{R}^n$ is a Gaussian white noise with stationary covariance $Q = \sigma^2 I$. Utilizing (9), where the expansion has been truncated to N-terms, the jth state equation in (10) can be written as

$$\dot{x}_N^{(j)}(t) = f^{(j)}(x_N(t), t) + \sqrt{\frac{2}{T}} \sigma \sum_{i=1}^{N} \zeta_i^{(j)}(\omega) \cos\left(\left(i - \frac{1}{2}\right) \frac{\pi t}{T}\right) \tag{11}$$

where $j = 1, \ldots, n$, and $x_N^{(j)}(t)$ is the jth state flow obtained with N-term noise KL approximation.

Notice that the Langevin form (10) can be written from (1), as long as the additive process noise in the SDE, is a second order stochastic process. For example, instead of the Wiener process, if the dynamics SDE has compound Poisson process, then the corresponding $\eta(\omega, t)$ in (10) would be Poisson white noise (see [26] for details). As $N \to \infty$, although the noise KL expansion converges to the true unstructured process noise in mean square sense, it is not apparent how the flow $x_N(t)$ is related to the SDE sample path $x(t)$. This is investigated next.

B. Convergence of Solution

Due to finite term truncation of the white noise KL exapansion, the solution of (11) is only an approximation of the SDE solution. Thus, it is imperative to quantify a notion of convergence, if any, between the solution of (11) and that of (10). In this section, we show that as $N \to \infty$, $x_N(t)$ converges to $x(t)$ in mean square sense, and then discuss about the rate of convergence.

1) Asymptotic Convergence:

Theorem 3: Let $x(\omega, t)$ be the solution of the nonlinear Itô SDE

$$dx(t) = f(x, t)dt + dW(\omega, t) \tag{12a}$$

$$\Rightarrow \frac{dx(t)}{dt} = f(x, t) + \eta(\omega, t), \tag{12b}$$

where $dW(\omega, t) = \eta(\omega, t) \, dt$, and $f : \mathbb{R}^n \times [0, T] \to \mathbb{R}^n$ satisfies the following:

1) non-explosion condition: $\exists \, D \geq 0$, s.t. $|f(x, t)| < D(1 + |x|)$ where $x \in \mathbb{R}^n$, $t \in [0, T]$;

2) Lipschitz condition: $\exists \, C \geq 0$, s.t. $|f(x, t) - f(\breve{x}, t)| < C |x - \breve{x}|$, where $x, \breve{x} \in \mathbb{R}^n$, $t \in [0, T]$.

Let $x_N(t)$ be solution of the ordinary differential equation (ODE), formed by N-term approximation of $\eta(\omega, t)$ in (12a), using orthonormal basis, which is given by,

$$\frac{dx_N(t)}{dt} = f(x_N, t) + \eta_N(\omega, t), \tag{13}$$

where $\eta_N(\omega, t)$ is the N-term approximation and $\mathbb{E}\left[\int_0^T \eta_N(\omega, t) dt\right] < \infty$. Then,

$$\lim_{N \to \infty} \mathbb{E}\left[|x(t) - x_N(t)|^2\right] = 0, \tag{14}$$

iff $x_N(t)$ is the KL expansion of $x(t)$.

Proof Throughout the proof, we will write $x(t)$ as $x(\omega, t)$ where $\omega \in \Omega$, the sample space of the random process $x(t)$. **First we will prove that** if (14) holds then $x_N(\omega, t)$ is the KL expansion of $x(\omega, t)$.

Let $\{\phi_m(t)\}_{m=1}^{\infty}$ be any orthonormal basis. Then $x(\omega, t)$ can be written as a convergent series sum

$$x(\omega, t) = \sum_{m=1}^{\infty} b_m c_m(\omega) \phi_m(t). \tag{15}$$

Let the solution of (13) be a $N-$term approximation of $x(\omega, t)$ which converge to true solution in mean square sense. Hence the truncation error e_N is given by

$$e_N(\omega, t) = \sum_{m=N+1}^{\infty} b_m c_m(\omega) \phi_m(t). \tag{16}$$

Projecting $x(\omega, t)$ in ϕ_m results

$$c_m(\omega) = \frac{1}{b_m} \int_0^T x(\omega, t) \phi_m(t) dt. \tag{17}$$

Hence

$$
\begin{aligned}
\mathbb{E}[e_N(\omega, t)] &= \sum_{m=N+1}^{\infty} \sum_{k=N+1}^{\infty} \phi_m(t) \phi_k(t) \int_0^T \int_0^T \mathbb{E}[x(\omega, t_1) x(\omega, t_2)] \phi_m(t_1) \phi_k(t_2) dt_1 dt_2 \\
&= \sum_{m=N+1}^{\infty} \sum_{k=N+1}^{\infty} \phi_m(t) \phi_k(t) \int_0^T \int_0^T C_{xx}(t_1, t_2) \phi_m(t_1) \phi_k(t_2) dt_1 dt_2,
\end{aligned}
\tag{18}
$$

where $C_{xx}(t_1, t_2)$ is the covariance function of $x(\omega, t)$. For convergence of error, the orthonormal basis $\phi_m(t)$ should minimize $\int_0^T \mathbb{E}[e_N(\omega, t)] dt$. Taking the orthonormality of $\phi_m(t)$ into account, we get

$$\int_0^T \mathbb{E}[e_N(\omega, t)] dt = \sum_{m=N+1}^{\infty} \int_0^T \int_0^T C_{xx}(t_1, t_2) \phi_m(t_1) \phi_m(t_2) dt_1 dt_2, \tag{19}$$

which is to be minimized over $\phi_m(t)$, subject to the constraint

$$\int_0^T \phi_m(t)\phi_k(t)dt = \delta_{mk} \ \forall m, k \in \mathbb{N}. \tag{20}$$

Introducing Lagrange multipliers b_m^2, our objective function $J(.)$ becomes

$$
\begin{aligned}
J(\phi_m(t)) &= \min_{\phi_m(t)} \sum_{m=N+1}^{\infty} \int_0^T \int_0^T C_{xx}(t_1, t_2)\phi_m(t_1)\phi_m(t_2)dt_1dt_2 \\
&\quad -b_m^2\left(\int_0^T \phi_m(t)\phi_m(t)dt - 1\right).
\end{aligned} \tag{21}
$$

Differentiating (21) with respect to $\phi_m(t)$ and setting the derivative to zero, we obtain

$$\int_0^T \left[\int_0^T C_{xx}(t_1, t_2)\phi_m(t_1)dt_1 - b_m^2\phi_m(t_2)\right]dt_2 = 0, \tag{22}$$

$$\Rightarrow \int_0^T C_{xx}(t_1, t_2)\phi_m(t_1)dt_1 = b_m^2\phi_m(t_2), \tag{23}$$

which is the Fredholm integral equation for the covariance function of random process $x(\omega, t)$. Hence $\phi_m(t)$ and b_m^2 are the eigenfunction-eigenvalue pairs of $C_{xx}(t_1, t_2)$. This completes our proof.

Now we will prove that, if $x_N(\omega, t)$ is the KL expansion of $x(\omega, t)$, then to ensure that solution of (13) converge to solution of (12a) in mean square sense, $x_N(\omega, t)$ must be the solution of (13). To prove this, we recall the following uniqueness conditions on solution of (12a) and KL expansion of a random process.

Proposition 1 ([31], Chap. 5): Given the non-explosion condition and the Lipschitz condition are satisfied for $f(\cdot, \cdot)$ in (12a). Let Z be a random variable, independent of the σ-algebra $\mathcal{F}_\infty^m$, generated by $\eta(\omega, t), t \geq 0$ and $\mathbb{E}\left[|Z|^2\right] \leq \infty$. Then the stochastic differential equation in (12a) where $t \in [0, T], X(\omega, 0) = Z$, has a unique t-continuous solution $x(\omega, t)$ adapted to the filtration $\mathcal{F}_t^Z$ generated by Z, and $\eta(\omega, t), t \geq 0$ and $\mathbb{E}\left[\int_0^T |x(\omega, t)|^2 dt\right] < \infty$.

Proposition 2 ([30], Chap. 2): The KL expansion of the random process $x(\omega, t)$, given by $x(\omega, t) = \sum_{m=1}^{\infty} \sqrt{\lambda_m}\xi_m(\omega)\phi_m(t)$, is unique.

Let us assume that $\breve{x}_N(\omega, t)$ is the KL expansion of $x(\omega, t)$. Furthermore, assume that $\breve{x}_N(\omega, t) \neq x_N(\omega, t)$, which is the solution of (13). Hence, we assume that although $\breve{x}_N(\omega, t)$ is the KL expansion of $x(\omega, t)$, it does not satisfy (13), whose solution converge to the solution of (12a) in mean square sense.

Proposition 1 and 2 tell us that the solution of SDE is unique and any random process has an unique KL expansion. Also (13) has unique solution as RHS of (13) satisfies Lipschitz condition. This can be proven as follows: for right hand side of (13) to satisfy Lipschitz condition, we must have,

$$|f(x, t) + \eta_N(\omega, t) - f(\breve{x}, t) - \eta_N(\omega, t)| \leq C|x - \breve{x}| \ \Rightarrow \ |f(x, t) - f(\breve{x}, t)| \leq C|x - \breve{x}|,$$

which is true since $f(\cdot,\cdot)$ itself satisfies Lipschitz condition. Hence (12a) has unique solution that admits a unique KL expansion. Also according to our assumption, the solution of (13) converge to solution of (12a) in mean square sense. This contradicts our assumption that $\breve{x}_N(\omega,t) \neq x_N(\omega,t)$, which completes the proof.

Theorem 3 states conditions upon the solutions of approximated and true systems for mean square convergence to hold, under certain assumptions on the nonlinearities. However, no condition is imposed on the initial states, which we investigate next.

Theorem 4: Given the stochastic dynamical system

$$dx(t) = f(x(t),t)dt + d\mathcal{W}(\omega,t), \tag{24}$$

and its corresponding Langevin form with N-term noise KL approximation:

$$dx_N^{(j)}(t) = f^{(j)}(x_N(t),t)\,dt + \sum_{i=1}^{N} \sqrt{\Lambda_i}\zeta_i^{(j)}(\omega)e_i(t)\,dt, \tag{25}$$

where $\lim\limits_{N\to\infty} \mathbb{E}\left|\mathcal{W}^{(j)}(\omega,t) - \sum\limits_{i=1}^{N} \sqrt{\Lambda_i}\zeta_i^{(j)}(\omega)e_i(t)\right|^2 = 0, \forall j = 1,2,\ldots,n$. If the drift vector field $f(\cdot,\cdot)$ is Lipschitz, and $x(0) = x_N(0)$, then $\lim\limits_{N\to\infty} \mathbb{E}|x(t) - x_N(t)|^2 = 0$.

Proof Integrating (24) and (25) and taking the expected value of square of the difference, we obtain

$$\mathbb{E}\left|x^{(j)}(t) - x_N^{(j)}(t)\right|^2 = \mathbb{E}\left[\left|(x^{(j)}(0) - x_N^{(j)}(0)) + \int_0^t \left(f^{(j)}(x,s) - f^{(j)}(x_N,s)\right)ds + \right.\right.$$

$$\left.\left.\int_0^t d(\mathcal{W}_s^{(j)} - \sum_{i=1}^{N} \sqrt{\Lambda_i}\zeta_i^{(j)}(\omega)e_i(s))\right|^2\right],$$

$$\leq \underbrace{\mathbb{E}\left|(x^{(j)}(0) - x_N^{(j)}(0))\right|^2}_{B_j \text{ (say)}} + \mathbb{E}\left|\int_0^t (f^{(j)}(x,s) - f^{(j)}(x_N,s))ds\right|^2 +$$

$$\mathbb{E}\left|\int_0^t d(W_s^{(j)} - \sum_{i=1}^{N} \sqrt{\Lambda_i}\zeta_i^{(j)}(\omega)e_i(s))\right|^2,$$

$$\leq B_j + t\mathbb{E}\int_0^t \left|f^{(j)}(x,s) - f^{(j)}(x_N,s)\right|^2 ds + \mathbb{E}\left|\int_0^t d(W_s^{(j)} - \sum_{i=1}^{N} \sqrt{\Lambda_i}\zeta_i^{(j)}(\omega)e_i(s))\right|^2, \tag{26}$$

where in the last step, we used Chebyshev's integral inequality. Consequently, we have

$$\lim_{N\to\infty} \mathbb{E}\left|x^{(j)}(t) - x_N^{(j)}(t)\right|^2 \le B_j + \lim_{N\to\infty} t\mathbb{E}\left[\int_0^t \left|f^{(j)}(x,s) - f^{(j)}(x_N,s)\right|^2 ds\right] +$$

$$\lim_{N\to\infty} \mathbb{E}\left|\int_0^t d(W_s^{(j)} - \sum_{i=1}^N \sqrt{\Lambda_i}\zeta_i^{(j)}(\omega)e_i(s))\right|^2. \tag{27}$$

Since the vector field f is Lipschitz, and the noise KL expansion is m.s. convergent, hence from (27) we get

$$\underbrace{\lim_{N\to\infty} \mathbb{E}\left|x^{(j)}(t) - x_N^{(j)}(t)\right|^2}_{v(t) \text{ (say)}} \le B_j + tC\int_0^t \lim_{N\to\infty} \mathbb{E}\left|x^{(j)}(s) - x_N^{(j)}(s)\right|^2 ds$$

$$\Rightarrow v(t) \le B_j + A\int_0^t v(s)ds \Rightarrow v(t) \le B_j \exp(At), \tag{28}$$

where the last step follows from Gronwall's inequality with $tC \le A, \forall t \in (0, T]$. Since $x(0) = x_N(0) \Rightarrow B_j = 0, \forall j = 1, \ldots, n$, therefore, $\lim_{N\to\infty} \mathbb{E}|x(t) - x_N(t)|^2 = 0$. ∎

Theorem 4 implies that, for any fixed initial condition, for sufficiently large N, solution of (25) converge to the solution of (24) in mean square sense. Hence, if the process noise term in Langevin equation is well-approximated using KL expansion, it entails a consistent approximation in the solution space.

Notice that if $x_N(0) \ne x(0)$, Theorem 4 may not hold. The following corollary quantifies when it holds for uncertain initial conditions.

Corollary 5: Suppose $x_N(0) \ne x(0)$. If $x_N(0)$ is the *generalized polynomial chaos* (gPC) expansion of $x(0)$, then $\lim_{N\to\infty} \mathbb{E}|x(t) - x_N(t)|^2 = 0$.

Proof If $x_N(0)$ is the gPC expansion of $x(0)$, then they converge in mean square sense [30]. Hence $\lim_{N\to\infty} \mathbb{E}[|x(0) - x_N(0)|^2] = 0$, which implies that $\lim_{N\to\infty} \mathbb{E}[|x(t) - x_N(t)|^2] = 0$, by using Gronwall's inequality. Hence the result. ∎

2) Rate of Convergence: It is known [32] that the rate-of-convergence of KL expansion is governed by the decay rate of the sum of the eigenvalues of the covariance function.

Example 1: (**Rate-of-convergence of KL expansion of Wiener process**) Consider the stochastic process $dx(\omega, t) = dW(\omega, t)$. Eigenvalues of the covariance function of $x(\omega, t)$ are given by Theorem 1. Thus, the truncation error due to N-term KL approximation of $x(\omega, t)$, is

$$\sum_{i=N+1}^\infty \lambda_i = \sum_{i=1}^\infty \lambda_i - \sum_{i=1}^N \lambda_i = \frac{4T}{\pi^2}\left[\frac{\pi^2}{8} - \frac{1}{8}\left(\pi^2 - 2\psi^{(1)}\left(N + \frac{1}{2}\right)\right)\right] = \frac{T}{\pi^2}\psi^{(1)}\left(N + \frac{1}{2}\right), \tag{29}$$
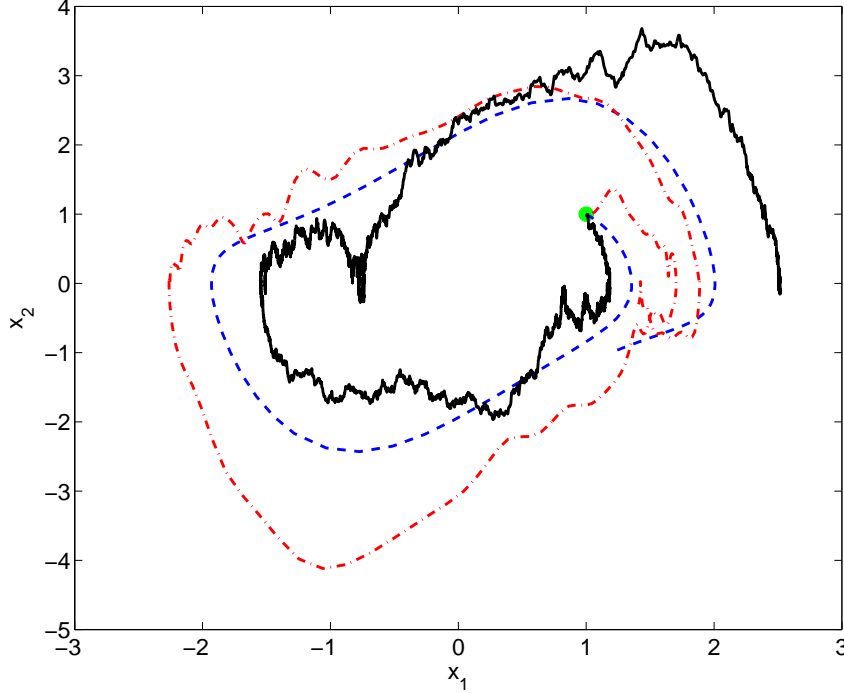
Fig. 1. This plot illustrates the asymptotic convergence results developed in Section II-B1 for Van der Pol oscillator given by (31). Starting from the same initial condition $(1, 1)$, denoted by the *filled circle*, the *dashed*, and *solid* curves show the deterministic (zero noise), and stochastic (SDE sample path with zero-mean additive Gaussian noise of variance 0.25) trajectories, respectively. The *dash-dotted* curve is the trajectory of the KL-approximated system with $N = 100$ terms, starting from the same initial condition, with process noise same as that of the SDE path. As N increases, the *dash-dotted* curve converges to the *solid curve* in mean-square sense.

which decays faster than $\mathcal{O}\left(N^{-1}\right)$, but slower than $\mathcal{O}\left(N^{-2}\right)$. Here $\psi^{(1)}\left(.\right)$ denotes the *trigamma function* [33].

Recently, it has been shown [34] that one can estimate the KL convergence rate by knowing the smoothness of the covariance function, even though the eigenvalues are not available analytically. The main result is that exponential convergence occurs when the covariance function is analytic. If the latter has Sobolev regularity, the KL expansion has algebraic decay. The following example on geometric Browinan motion (GBM) illustrates this.

Example 2: (**Rate-of-convergence of KL expansion of GBM**) Consider the GBM $dx\left(\omega, t\right) = ax\left(\omega, t\right)dt + bx\left(\omega, t\right)dW\left(\omega, t\right)$, where a, b are deterministic constants. The covariance function can be computed from (5) as $C\left(t_1, t_2\right) = \left(x_0\left(\omega\right)\right)^2 e^{a(t_1+t_2)}\left(e^{b^2 t_2} - 1\right)$. Since $C\left(t_1, t_2\right)$ is analytic, the convergence is exponential.

In the nonlinear estimation setting, the covariance function of the process being estimated, is not known beforehand. What is known here is the dynamics $f\left(.\right)$. This motivates us to investigate the connection

between the regularity of $f(.)$ with the regularity of $C(t_1, t_2)$. However, such an analysis has not been done here and will be a topic of our future work.

III. KLPF FORMULATION

In this section, we propose a methodology to predict the evolution of uncertainty in KL-approximated system using the PF operator. Let us consider the KL approximated dynamics (11) where the state $x_N(t) \in \mathbb{R}^n$ has an initial PDF $\varphi_0 \triangleq \varphi^-(x_N(0), 0)$. We next apply MOC to PF operator, using the formulation given in [21], and form an augmented dynamical system in $(n+1)$ dimensions, given by

$$\dot{x}_N^{(j)}(t) = f^{(j)}(x_N(t)) + \sqrt{\frac{2}{T}} \sigma \sum_{i=1}^{N} \zeta_i^{(j)}(\omega) \cos\left(\left(i - \frac{1}{2}\right) \frac{\pi t}{T}\right), \tag{30a}$$

$$\dot{\varphi}_N^- = -\operatorname{div}\left(f(x_N(t), t)\right) \varphi_N^-, \tag{30b}$$

where $j = 1, 2, \ldots, n$; $\operatorname{div}\left(f(x_N(t), t)\right) \triangleq \sum_{j=1}^{n} \frac{\partial f^{(j)}}{\partial x_N^{(j)}}$, and $\varphi_N^- \triangleq \varphi^-(x_N(t), t)$. Equation (30) describes the prior dynamics along the approximated state flow $x_N(t)$. Since $x_N(t) \overset{\text{m.s.}}{\to} x(t)$ as $N \to \infty$ (Theorem 4), hence $\lim_{N \to \infty} \varphi_N^- = \varphi^-(x(t), t)$. In other words, (30) provides a way to approximate the true prior φ^- that satisfies (3), while obviating the computational nuisances of solving a second order PDE. Details of the MOC solution framework can be found in [7], [23].

A. Illustrative Examples

The proposed methodology is applied to a Vanderpol's oscillator, given by

$$\ddot{x}(t) = \left(1 - x^2(t)\right) \dot{x}(t) - x(t) + \eta(\omega, t), \tag{31}$$

and to a Duffing oscillator, with dynamics,

$$\ddot{x}(t) = 10x(t) - 30x^3(t) - 10\dot{x}(t) + \eta(\omega, t), \tag{32}$$

with $\eta(\omega, t)$ having autocorrelation $2\pi \mathbf{I}$. The initial state uncertainty, has a PDF $\rho_0 \sim \mathcal{N}\left([0, 0]^T, \operatorname{diag}(1, 1)\right)$ for both the systems in (31) and (32).

Taking $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$, the augmented dynamical system for the Vanderpol's oscillator, is given by

$$\dot{x}_1(t) = x_2(t), \tag{33a}$$

$$\dot{x}_2(t) = \left(1 - x_1^2(t)\right) x_2(t) - x_1(t) + \eta(\omega, t), \tag{33b}$$

$$\dot{\rho}(x(t)) = -\left(1 - x_1^2(t)\right) \rho(x(t)), \tag{33c}$$

and for the Duffing oscillator, is given by

$$\dot{x}_1(t) = x_2(t), \tag{34a}$$

$$\dot{x}_2(t) = 10x_1(t) - 30x_1^3(t) - 10x_2(t) + \eta(\omega, t), \tag{34b}$$

$$\dot{\rho}(x(t)) = 10\rho(x(t)). \tag{34c}$$

Next, the initial PDF ρ_0 is sampled with sample size of $\nu = 5000$. For the Vanderpol's oscillator, final
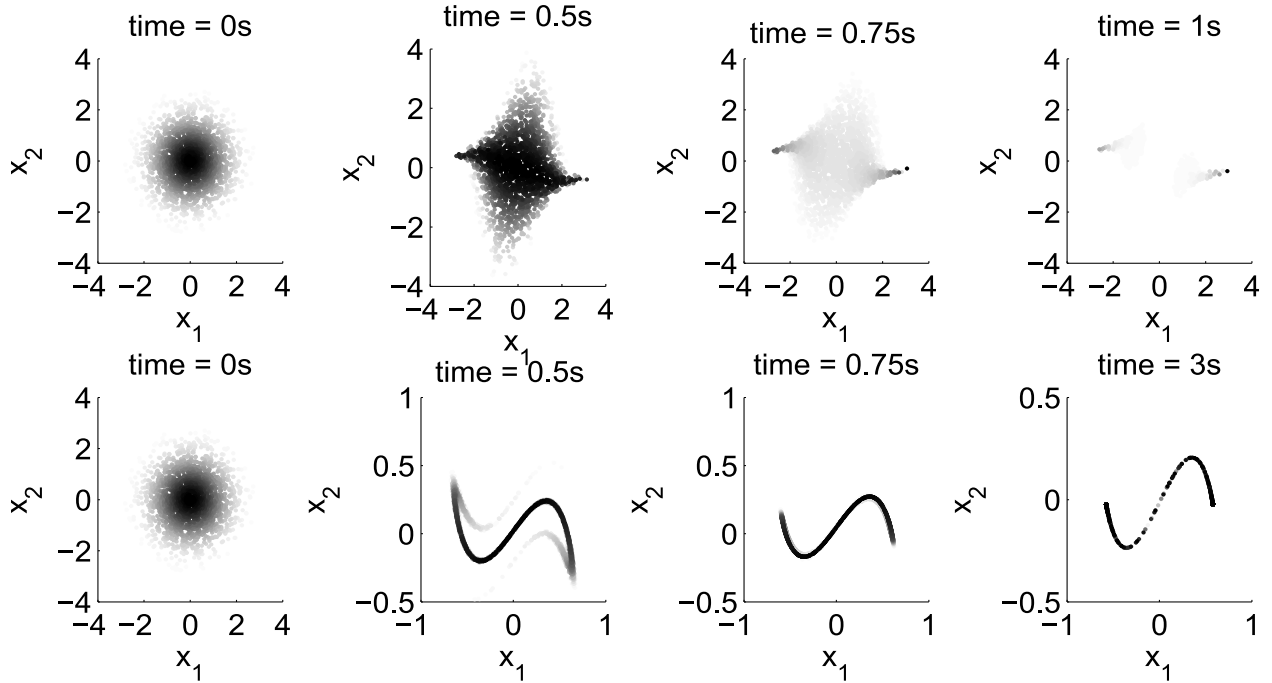

Fig. 2. Uncertainty Propagation for Vanderpol's oscillator (31) in top row, and Duffing oscillator (32) in bottom row. Darker regions have higher PDF value than lighter regions.

time T is taken to be $1s$, for the Duffing oscillator $T = 3s$. Number of terms in the KL expansion is kept fixed as $N = 7$. Figure 2 shows the evolution of probability densities with time for the two oscillators. It is observed that, for the Vanderpol's oscillator the probability mass accumulates along the limit cycle and for the Duffing oscillator, we get a bimodal PDF at final time. This is in agreement with the physical intuition for behavior of these systems.

B. Extensions

It is important to recognize that the main idea behind KLPF is to derive an approximate *ODE initial value problem (IVP)* from the original stochastic description, thereby facilitating an MOC based PF

operator computation in exact arithmetic. Thus, it is possible to extend our KLPF formulation for more general class of stochastic systems, not necessarily described by an SDE IVP, as long as we can arrive at some ODE IVP approximation. In the SDE IVP case, this has been achieved through a two-step "first KL, then PF" algorithm (Figure 3). For general stochastic evolution equations, an intermediate approximation step is needed between the KL approximation and MOC based PF computation step.
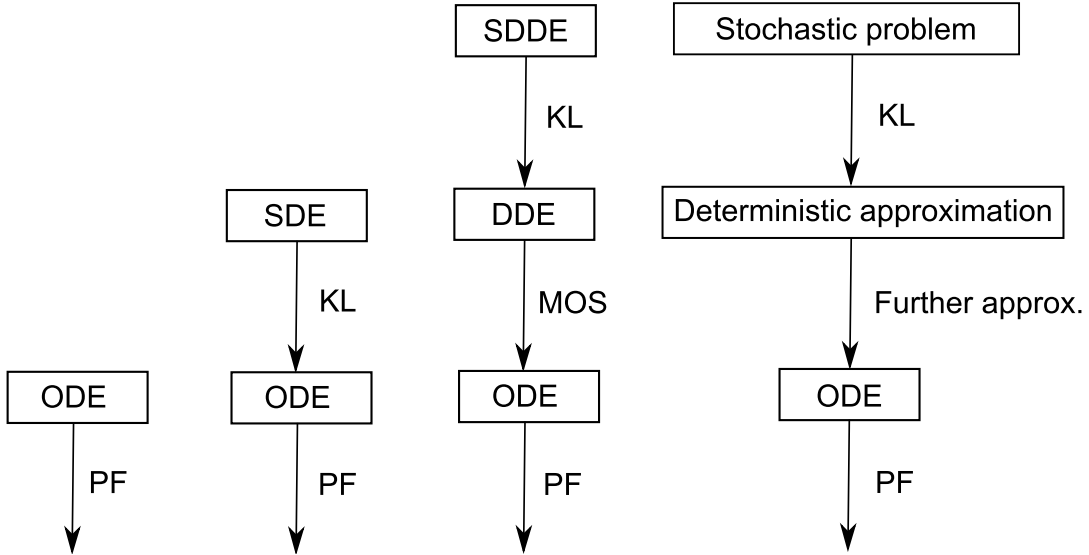

Fig. 3.   Schematic of the generalized KLPF formulation. As the right-most flow-chart shows, an intermediate approximation step is needed for generic stochastic evolution, between KL approximation and PF operator based computation.

1) Example: Uncertainty quantification for stochastic delay differential equation (SDDE): Consider a dynamics described by nonlinear SDDE [35]

$$
\begin{aligned}
dx\,(t) &= f\,(t, x\,(t)\,, x\,(t - \tau))\,dt + dW\,(\omega, t)\,, \qquad \forall\,t > 0,\\
x\,(t) &= \psi\,(t)\,, \qquad \forall\,t \in (-\tau, 0],
\end{aligned}
\tag{35}
$$

where τ denotes a fixed time-delay, $\psi\,(t)$ is the history function, and $W\,(\omega, t)$ denotes the Wiener process. The drift vector field depends on both the current state $x\,(t)$, and the delayed state $x\,(t - \tau)$. Given (35), solving the *delay Fokker-Planck equation* [36] is computationally hard for general nonlinear drift. Instead, we can write the KL approximated dynamics corresponding to (35) as

$$
\begin{aligned}
\frac{dx_N}{dt} &= f\,(t, x_N\,(t)\,, x_N\,(t - \tau)) + \eta_N\,(\omega, t)\,, \qquad \forall\,t > 0,\\
x_N\,(t) &= \psi\,(t)\,, \qquad \forall\,t \in (-\tau, 0],
\end{aligned}
\tag{36}
$$

which is a DDE boundary value problem (BVP), with $\eta_N(\omega, t)$ being the N-term KL approximation of Gaussian white noise. We next use *method-of-steps*[1] (MOS) [37] to further reduce this infinite-dimensional BVP to a sequence of finite-dimensional ODE IVPs, as shown:

$$
\begin{aligned}
\dot{x}_N^{(1)}(t) &= f\left(t, x_N^{(1)}(t), \psi(t-\tau)\right) + \eta_N(\omega, t), & \forall\, t \in (0, \tau], \\
\dot{x}_N^{(2)}(t) &= f\left(t, x_N^{(2)}(t), x_N^{(1)}(t-\tau)\right) + \eta_N(\omega, t), & \forall\, t \in (\tau, 2\tau], \\
&\vdots \\
\dot{x}_N^{(m)}(t) &= f\left(t, x_N^{(m)}(t), x_N^{(m-1)}(t-\tau)\right) + \eta_N(\omega, t), & \forall\, t \in ((m-1)\tau, m\tau], \quad (37)
\end{aligned}
$$

where $m \in \mathbb{Z}^+$. Now we can apply the MOC based PF formulation to (37) by solving the following pair of ODEs in the jth interval $((j-1)\tau, j\tau]$:

$$
\begin{aligned}
\dot{x}_N^{(j)}(t) &= f\left(t, x_N^{(j)}(t), x_N^{(j-1)}(t-\tau)\right) + \eta_N(\omega, t), & \text{(38a)} \\
\dot{\rho}^{(j)} &= -\text{div}\left(f\left(t, x_N^{(j)}(t), x_N^{(j-1)}(t-\tau)\right)\right)\rho^{(j)}, & \text{(38b)}
\end{aligned}
$$

with the initial conditions being the respective terminal values of the $(j-1)^{\text{th}}$ interval. Eqn. (38) computes the evolution of joint PDF $\rho^{(j)}\left(t, x_N^{(j)}(t)\right)$ along the trajectory $x_N^{(j)}(t)$.

IV. Nonlinear Filtering Using KLPF Method

Let us consider a nonlinear dynamical system with states $x \in \mathbb{R}^n$ being estimated from a continuous-time noisy dynamics model, and a discrete-time nonlinear sensor model $h(\cdot, \cdot)$ with noisy measurements $\tilde{y}_k \in \mathbb{R}^m$, given by

$$
\begin{aligned}
\dot{x}(t) &= f(x(t), t) + \eta(\omega, t), & \text{(39a)} \\
\tilde{y}_k &= h(x_k, t_k) + \vartheta_k, & \text{(39b)}
\end{aligned}
$$

where $\{t_k\}_{k=0}^r$ are time instances when measurements are obtained, and $\{x_k\}_{k=0}^r$ are the corresponding states. $\eta(\omega, t)$ and ϑ_k are zero mean delta correlated Gaussian random processes with covariances Q and R, respectively. Let $\varphi_k^-(.)$ and $\varphi_k^+(.)$ denote the prior and posterior PDFs at time t_k. The state estimation algorithm described next, is developed in similar manner as [23].

---

[1]MOS provides the intermediate approximation step from DDE BVP to ODE IVP, as shown in the third column of Figure 3.

A. Step 1: Initialization

We sample ν particles from the support of the initial state PDF φ_0^-. Let the particles be $\{x_{0,i}\}_{i=1}^\nu$ with corresponding PDF values $\{\varphi_{0,i}^-\}_{i=1}^\nu$. The following steps are then performed sequentially starting from $k = 1$.

B. Step 2: Propagation

Consider the N-term noise KL approximated version of (39a), and the corresponding MOC based PF operator equation, given together by (30). With the initial states at $(k-1)^{\text{th}}$ step as $[x_{k-1,i}, \varphi_{k-1,i}]$, (30) is integrated over the interval $[t_{k-1}, t_k]$ for each particle, to get the states $x_{k,i}$ and the prior PDF $\varphi_{k,i}^-$ at t_k.

C. Step 3: Update

Using the Gaussian nature of measurement noise, for each particle, the likelihood function is determined as

$$\varrho_{k,i}(\tilde{y}_k|x_k = x_{k,i}) = \frac{1}{\sqrt{(2\pi)^m |R|}} \times \exp -\frac{1}{2}\left[(\tilde{y}_k - h(x_{k,i}, t_k))^\top R^{-1} (\tilde{y}_k - h(x_{k,i}, t_k)) \right], \qquad (40)$$

where $|R|$ denotes the determinant of the measurement noise covariance matrix. The posterior PDF value for the ith particle, can then be obtained as

$$\varphi_{k,i}^+ = \frac{\varrho_{k,i}(\tilde{y}_k|x_k = x_{k,i})\varphi_{k,i}^-}{\displaystyle\sum_{i=1}^\nu \varrho_{k,i}(\tilde{y}_k|x_k = x_{k,i})\varphi_{k,i}^-}. \qquad (41)$$

The true posterior φ^+ is the limit of the above expression as $\nu \to \infty$. However, unlike particle filters, here we compute the prior density explicitly using PF operator. Hence, accounting the dynamics of the prior itself, makes (41) an accurate representation of the posterior PDF, as no histogram approximation is needed here.

D. Step 4: Getting the State Estimate

Computation of the state estimate for the kth step, depends on the desired optimality criterion involving φ^+. Most commonly used criteria are *maximum likelihood criterion* (MLC), *minimum error criterion* (MEC), and *minimum variance criterion* (MVC) [38]. For the examples following this section, we have used MVC to obtain the state estimate, which is given by $\hat{x}_k = \arg\min_x \displaystyle\sum_{i=1}^\nu \|x - x_{k,i}\|^2 \varphi_k^+(x_{k,i}) = \displaystyle\sum_{i=1}^\nu x_{k,i}\, \varphi_k^+(x_{k,i})$.

In principle, one can use additional resampling techniques like particle filtering, to avoid degeneracy and increase diversity amongst particles [15]. However, qualitatively, since histogram approximation was not used for prior PDF construction, the proposed method is less sensitive to the issue of degeneracy.

V. Numerical Results

In this section, we compare the results obtained using KLPF filter with those obtained using particle filter. The particle filter used here is the *sequential importance resampling* particle filter, where the threshold particle number is optimized to get the best performance (Chapter 3, Section 3.5.1 of ref. [15]). Moreover, for particle filter implementation, the SDE is integrated using *Euler-Maruyama* scheme [39].

A. Comparison of Posterior Densities of Particle Filter and KLPF Estimator

We compare the KLPF estimator to particle filters, by investigating the posterior PDFs obtained from the two methods after each filtering step. Our objective is to illustrate the performance improvement achieved by KLPF compared to particle filter, for cases where the estimation problems are known to admit finite dimensional sufficient statistics. In such cases, knowledge of the true posterior enables us to conclude the performance improvement accomplished by KLPF.

1) Case I: Kalman Filter: For a linear system with Gaussian uncertainty, we know the exact evolution of φ^+ through Kalman filter. We use the Wasserstein metric W [40], [41], [42] to compute the distance between the algorithm predicted posterior and the reference posterior, i.e. $W(\varphi^+_{\text{Particle}}, \varphi^+_{\text{Kalman}})$ and $W(\varphi^+_{\text{KLPF}}, \varphi^+_{\text{Kalman}})$. The Wasserstein metric computes the distance between the shapes of the statistical distributions under comparison, and can be interpreted as the minimum amount of work needed to morph one to the other. If $\varsigma_1 \sim \mathcal{N}(\mu_1, \Sigma_1)$ and $\varsigma_2 \sim \mathcal{N}(\mu_2, \Sigma_2)$, then the Wasserstein metric admits a closed form expression:

$$W(\varsigma_1, \varsigma_2) = \sqrt{||\mu_1 - \mu_2||_2^2 + \text{tr}(\Sigma_1 + \Sigma_2 - 2(\sqrt{\Sigma_1}\Sigma_2\sqrt{\Sigma_1})^{\frac{1}{2}})}. \tag{42}$$

In general, W between two arbitrary PDFs φ and $\widehat{\varphi}$, is defined as

$$W(\varphi, \widehat{\varphi}) \triangleq \left(\inf_{\rho \in \mathcal{M}(\varphi, \widehat{\varphi})} \mathbb{E}\left[\| x - \widehat{x} \|_2^2 \right] \right)^{1/2}, \tag{43}$$

where x and $\widehat{x}$ are the random vectors corresponding to the PDFs φ and $\widehat{\varphi}$, respectively, and $\mathcal{M}(\varphi, \widehat{\varphi})$ is the space of all joint PDFs with first marginal φ, and second marginal $\widehat{\varphi}$. It is known [43] that (43) amounts to solving an infinite dimensional linear program (LP).

As a concrete numerical example, let us consider the state dynamics

$$\dot{x}(t) = -0.05I_2\ x(t) + [1 \quad 1]^\top \eta(t), \tag{44}$$

where the state $x(t) \in \mathbb{R}^2$, and the identity matrix $I_2 \in \mathbb{R}^{2 \times 2}$. Let the discrete measurement model be given by

$$\tilde{y}_k = \begin{bmatrix} 1 & 1 \end{bmatrix} x_k + v_k, \tag{45}$$

where $\eta(t)$ and v_k are assumed to be zero mean delta correlated Gaussian noise processes with variances $Q = 1/8$ and $R = 1/4$, respectively. Further, let $\varphi_0 = \mathcal{N}([1, 1], \text{diag}(1, 1))$.

We draw 100 samples, each of sample size 500 from φ_0, and plot (Figure 4) the means and standard deviations of $W(\varphi^+_{\text{Particle}}, \varphi^+_{\text{Kalman}})$ and $W(\varphi^+_{\text{KLPF}}, \varphi^+_{\text{Kalman}})$, computed from (42). This plot shows that the KLPF filter posterior remains indeed closer to the Kalman posterior, compared to the particle filter posterior.
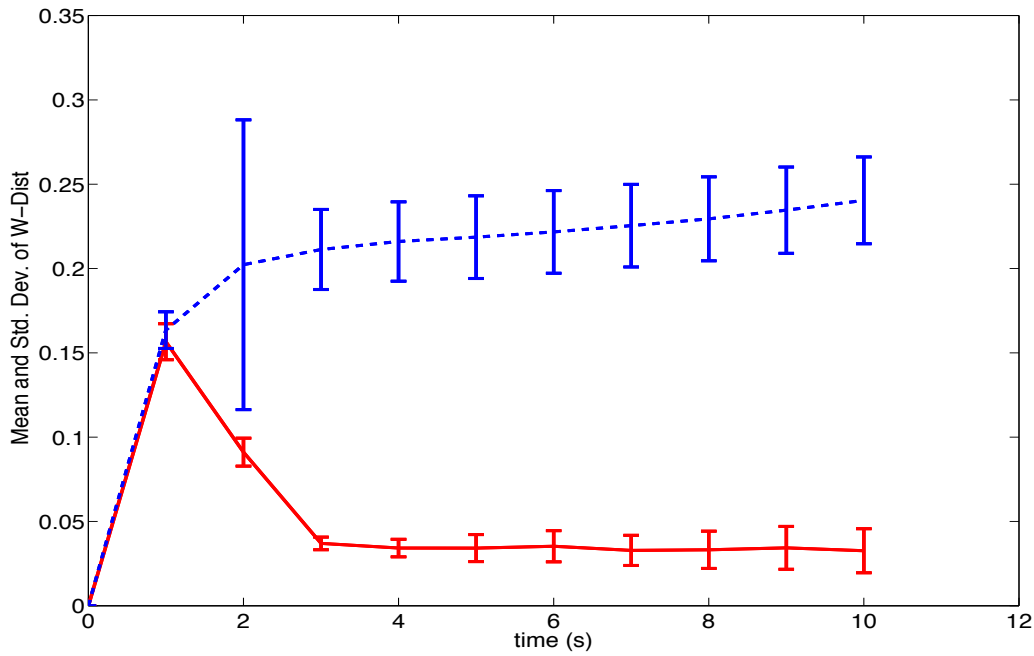

Fig. 4. Plot of mean and standard deviation of the Wasserstein distances of the KLPF estimator (*solid line*) and the particle filter (*hyphenated line*) from the Kalman filter. The vertical lines about the means represent $\pm 1\sigma$ limits.

2) Case II: Beneš Filter: Beneš filter is one of the few nonlinear filters which admit a known finite-dimensional solution of the nonlinear estimation problem. Here, the nonlinear drift is assumed to satisfy a Riccati differential equation [44] and the measurement model is taken to be affine in states. For simplicity,

we consider the continuous-time scalar Beneš filtering problem of the form:

$$\begin{aligned} dx\left(t\right) &= \frac{\kappa e^{x} - e^{-x}}{\kappa e^{x} + e^{-x}} \, dt + d\mathcal{W}\left(\omega, t\right), \\ dy\left(t\right) &= x\left(t\right) \, dt + d\mathcal{V}\left(\omega, t\right), \end{aligned} \tag{46}$$

with $\kappa = 0.5$ and deterministic initial condition x_0. The process and measurement noise densities are $\mathcal{N}\left(0, Q\right)$ and $\mathcal{N}\left(0, R\right)$ respectively, with $Q = 1$, $R = 10$. It can be shown [45] that the drift nonlinearity in (46) satisfy the necessary Riccati condition and the resulting solution [46] is given by the normalized posterior density

$$\varphi\left(x\left(t\right) | \mathcal{Y}_t\right) = \sqrt{\frac{\coth(t)}{2\pi}} \times \left(\frac{\kappa e^{x} + e^{-x}}{\kappa e^{I_t(y(\omega))} + e^{-I_t(y(\omega))}} \right) \exp\left(-\frac{1}{2} \Gamma\left(t\right) \right), \tag{47}$$

where $\mathcal{Y}_t$ is the history (filtration) till time t, and

$$\begin{aligned} I_t\left(y\left(\omega\right)\right) &:= \operatorname{sech}(t) \left[x_0 + \int_0^t \sinh(s) dy_s\left(\omega\right) \right], \\ \Gamma\left(t\right) &:= \tanh(t) + \coth(t) \left(x - I_t\left(y\left(\omega\right)\right)\right)^2. \end{aligned} \tag{48}$$

Notice that for this nonlinear non-Gaussian estimation problem, unlike Kalman filter case, we can not write the Wasserstein distance between the true posterior (47) and particle filter/KLPF posterior, as an analytical expression in terms of the respective sufficient statistics. Thus, in order to compute the Wasserstein time history, we resort to (43). At each time, we sample (47) using the Metropolis-Hastings MCMC technique [47], and solve the LP between the sampled true Beneš posterior and particle filter/KLPF posterior, to result the normalized Wasserstein trajectories shown in Figure 5. Like the Kalman filter case, as time progresses, KLPF posterior gets closer, compared to particle filter, to true Beneš posterior.
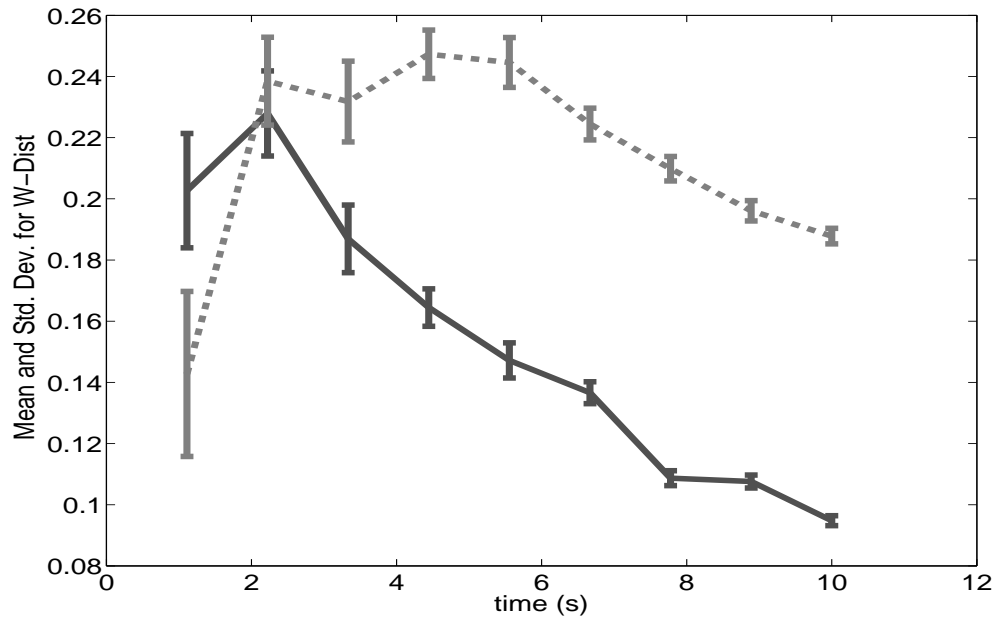
Fig. 5. Plot of mean and standard deviation of the Wasserstein distances of the KLPF estimator (*solid line*) and the particle filter (*hyphenated line*) for the Beneš filter. The vertical lines about the means represent $\pm 1\sigma$ limits.

B. Application to Hypersonic Reentry

The KLPF filtering technique is applied next to estimate states of a hypersonic spacecraft entering the atmosphere of Mars. The entry dynamics is given by the "noisy" version of Vinh's equation [48]:

$$\dot{r} = v \sin\gamma + \eta_r, \tag{49a}$$

$$\dot{\theta} = \frac{v \cos\gamma \cos\xi}{r \cos\lambda} + \eta_\theta, \tag{49b}$$

$$\dot{\lambda} = \frac{v \cos\gamma \sin\xi}{r} + \eta_\lambda \tag{49c}$$

$$\dot{v} = -\frac{\rho v^2}{2 B_c} - g \sin\gamma - \overline{\Omega}^2 r \cos\lambda (\sin\gamma \cos\lambda - \cos\gamma \sin\lambda \sin\xi)$$
$$+ \eta_v \tag{49d}$$

$$\dot{\gamma} = \left(\frac{v}{r} - \frac{g}{v} \right) \cos(\gamma) + \frac{\rho}{2 B_c} \left(\frac{L}{D} \right) v \cos\sigma + 2\overline{\Omega} \cos\lambda \cos\xi +$$
$$\frac{\overline{\Omega}^2 r}{v} \cos\lambda (\cos\gamma \cos\lambda + \sin\gamma \sin\lambda \sin\xi) + \eta_\gamma \tag{49e}$$

$$\dot{\xi} = \frac{\rho}{2 B_c} \left(\frac{L}{D} \right) v \sin\sigma - \frac{v}{r} \cos\gamma \cos\xi \tan\lambda + 2\overline{\Omega}(\tan\gamma \cos\lambda \sin\xi$$
$$- \sin\lambda) - \frac{\overline{\Omega}^2 r}{v \cos\gamma} \sin\lambda \cos\lambda \cos\xi + \eta_\xi, \tag{49f}$$

where r is the distance of the vehicle from the planet's center, θ and λ are Mars-centric longitude and latitude respectively, v is the total velocity of the vehicle, γ is the flight path angle and ξ is the azimuth angle. The symbol g denotes the acceleration due to gravity, B_c denotes the ballistic coefficient, $\frac{L}{D}$ stands for the lift-to-drag ratio of the the vehicle, ρ is the atmospheric density given by $\rho = \rho_0 e^{\left(\frac{h_2 - h}{h_1} \right)}$, where ρ_0, h_1 and h_2 are constants depending on the planet's atmospheric model, $h = r - R_m$ is the height measured from the planet's surface and R_m is the mean equatorial radius of the planet. η_r, η_θ, η_λ, η_v, η_γ and η_ξ are zero mean, delta correlated Gaussian white noises. Description of the constants in (49a) through (49f), used to simulate reentry in Martian atmosphere, are given in Table I [23].

The system is assumed to have *initial Gaussian state uncertainty* with mean vector and covariance matrix given by,

$$\mu_0 = [R_m + 54 \text{ Km}, -60^o, 30^o, 2.4 \text{ Km/s}, -9^o, 0.0573^o]^\top,$$

and

$$\Sigma_0 = \text{diag} \left(5.4 \text{ Km}, 3^o, 3^o, 240 \text{ m/s}, 0.9^o, 0.0057^o \right).$$

The discrete measurement model consists of six measurements, and is given by

$$\tilde{\mathbf{y}} = [\bar{q}, H, \gamma, \theta, \lambda, \xi]^\top + \vartheta \tag{50}$$

TABLE I

PARAMETER VALUES FOR (49)

Description of Constants	Value
Radius of Mars	$R_m = 3397 \times 10^3$ m
Acceleration due to gravity of Mars	$g = 3.71$ m/s^2
Ballistic coefficient of the vehicle	$B_c = 72.8$ kg/m^2
Lift-to-drag ratio of the vehicle	$\frac{L}{D} = 0.3$
Density at the surface of mars	$\rho_0 = 0.0019$ kg/m^3
Scale height for density computation	$h_1 = 9.8$ km $h_2 = 20$ km
Escape velocity of Mars	$v_c = 5.027$ km/s
Angular velocity of Mars	$\overline{\Omega} = 7.0882 \times 10^{-5}$ rad/s
Bank angle	$\sigma = 0^o$

where $\bar{q} = 1/2\rho v^2$ is the dynamic pressure and $H = K\rho^{\frac{1}{2}}v^{3.15}$ is the heating rate, with $K = 4.47228 \times 10^{-9}$ being the scaled material heating coefficient [49]. $\vartheta \in \mathbb{R}^6$ is a vector of zero mean delta correlated discrete measurement noise. For the present case, measurements update interval was fixed to 20s.

Equation (49) and (50) are non-dimensionalized before applying the filtering algorithms. The characteristic mass, length and time, for this non-dimensionalization, are mass of the reentry vehicle (2800 kg), R_m and $R_m/v_c = 675.7$ s, respectively. In scaled units the process noises are assumed to have covariance of $Q = 3.6 \times 10^{-5}$, and the measurement noise has covariance of $R = 3.6 \times 10^{-3} \times \mathcal{I}_6$, where $\mathcal{I}_n$ is $n \times n$ identity matrix. For the KLPF based estimator, the number of terms used in the KL expansion of each process noises in (49f) is fixed to 10.

Figure 6 shows the plots for square root of the difference between variance and Cramer-Rao lower bounds (CRLBs) for each state. The CRLBs are calculated after the states in (49f) are discretized for numerical integration. We use the formula given in Ref. [15] (Chapter 4) to calculate the CRLBs. If x_k is the discretized state at a given time t_k, then the CRLBs are given by J_k which is the Schur complement of the transpose of *Fisher's information matrix*. The J_k is obtained recursively by solving the following equation,

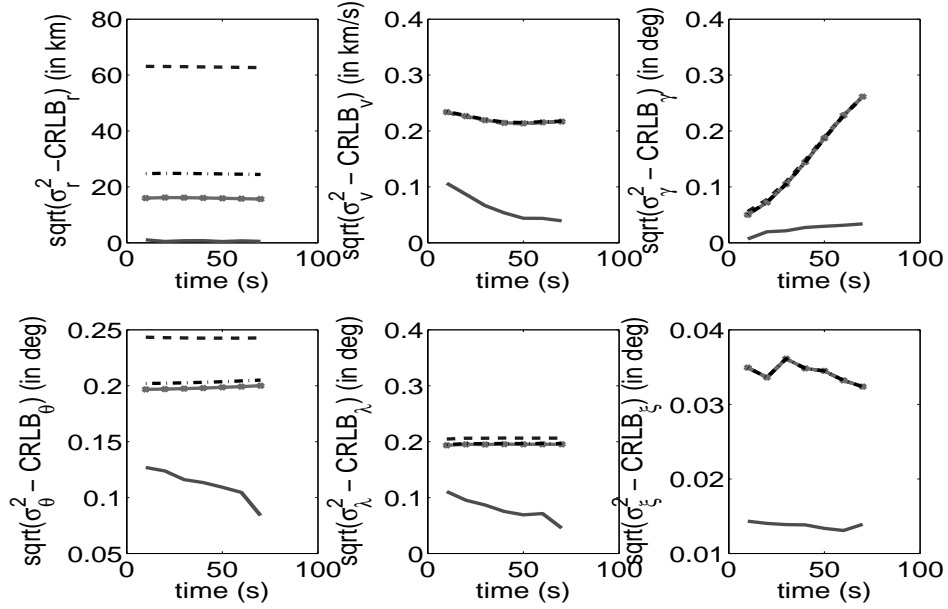$$J_{k+1} = D_k^{22} - D_k^{21}(J_k + D_k^{11})^{-1}D_k^{21}$$

Fig. 6. Plots for $\sqrt{\sigma^2 - CRLB}$ for all the six states. The solid line represents KLPF filter (3000 particles). The hyphenated, hyphen-dotted and solid-asterixed lines represent particle filters with 3000, 20000 and 50000 particles respectively.

where,

$$D_k^{11} = -\mathbb{E}\left[\nabla_{x_k}[\nabla_{x_k} \log \varphi_{k+1}^-]^\top\right],$$

$$D_k^{21} = -\mathbb{E}\left[\nabla_{x_k}[\nabla_{x_{k+1}} \log \varphi_{k+1}^-]^\top\right],$$

$$D_k^{22} = -\mathbb{E}\left[\nabla_{x_{k+1}}[\nabla_{x_{k+1}} \log \varphi_{k+1}^-]^\top\right]$$
$$- \mathbb{E}\left[\nabla_{x_{k+1}}[\nabla_{x_{k+1}} \log \varrho(\tilde{y}_{k+1}|x_{k+1})]^\top\right],$$

where, φ_{k+1}^- and $\varrho(\tilde{y}_{k+1}|x_{k+1})$ are the prior PDF of the states and the likelihood function at time t_{k+1}, respectively. It can be seen that for a additive Gaussian noise the above equations have a simple form.

The KLPF estimator with sample size 3000 is compared with particle filters with 3000, 20000 and 50000 particles. We can see that the difference for KLPF estimator is lower than that of the particle filters for all the states. Hence, for the present case, the solution obtained from proposed estimation scheme is closer to the lowest possible minimum variance solution than that obtained from particle filters.

Next, we plot the final posterior univariate and bivariate marginals obtained from the state estimation algorithms. The marginal PDFs were calculated using the algorithm given in ref. [21]. Figure 7 shows
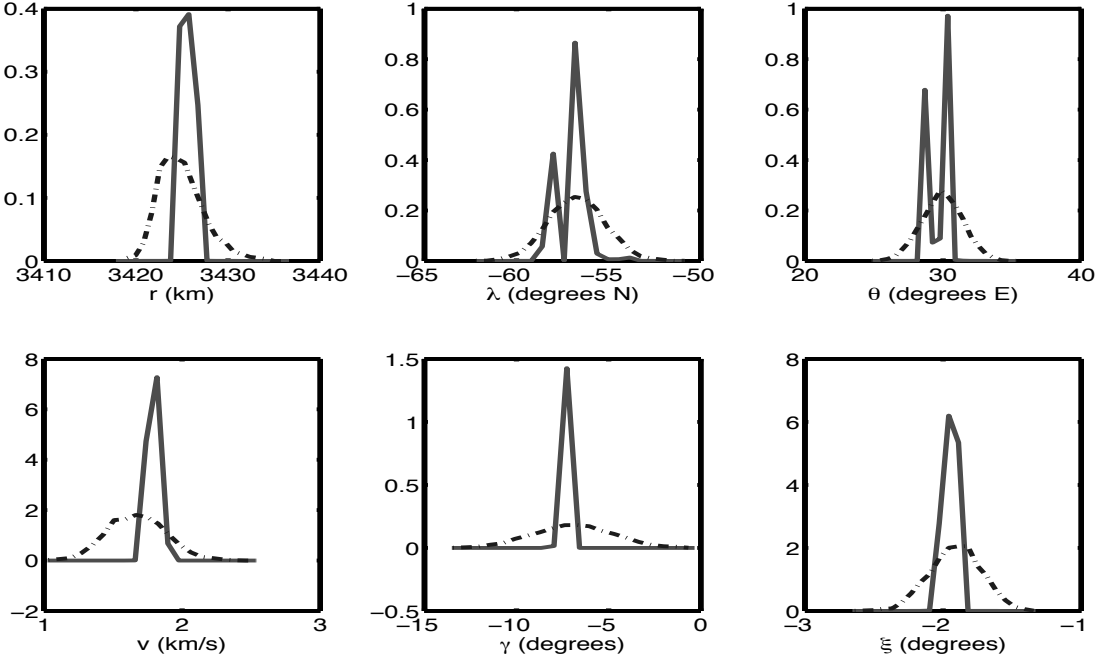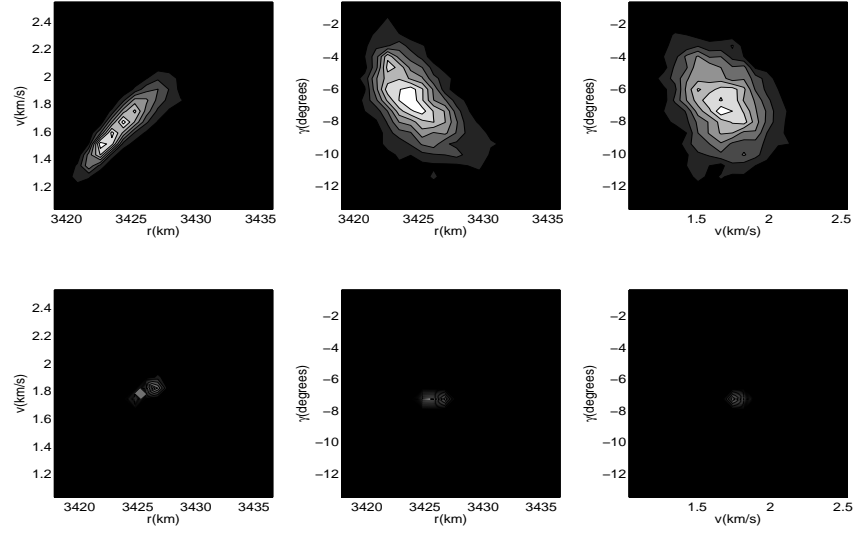
Fig. 7. Final posterior unidimensional marginal PDF for all states obtained from KLPF estimator (solid line) and particle filter (hyphenated line) with 3000 particles.

comparison of univariate marginals obtained from KLPF estimator and particle filter, both with 3000 particles. Figure 8 plots bivariate marginal PDFs obtained from KLPF estimator and particle filter. It can be observed that the KLPF estimator is able to reduce variance and capture localization of uncertainty better than the particle filter with same number of particles.

The KLPF estimator performs better than the particle filter with larger number of particles. Figure 9 and Fig. 10 show plots for the final posterior univariate and bivariate PDFs respectively for KLPF estimator with 3000 particles and particle filter with 50000 particles. It can be seen that, the KLPF estimator reduces variance and captures uncertainty localization better than the particle filter, although the particle filter has larger number of particles than the previous case.

VI. CONCLUSION

This work presents a method of approximating the noise term in stochastic dynamical systems using KL expansion. It is proved that the solution of the resulting approximate dynamical system converge to the true solution in mean square sense. A technique to propagate uncertainty in SDEs combining the above mentioned method and PF operator is proposed. A nonlinear estimation technique is developed using

(a) Top row- particle filter, bottom row- KLPF estimator.
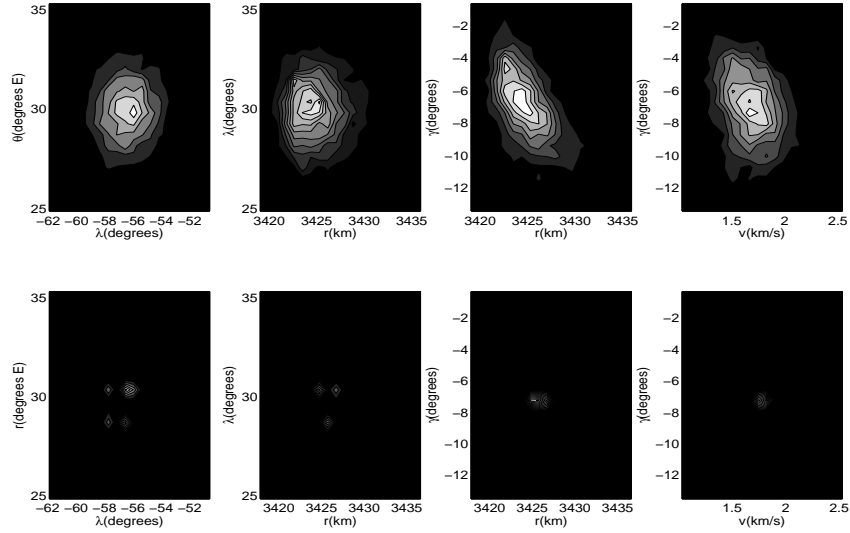
(b) Top row- particle filter, bottom row- KLPF estimator.

Fig. 8. Plots for the final posterior bivariate marginal PDFs obtained from KLPF estimator and particle filter with 3000 particles. The darker regions represent lower PDF value and the lighter regions higher.
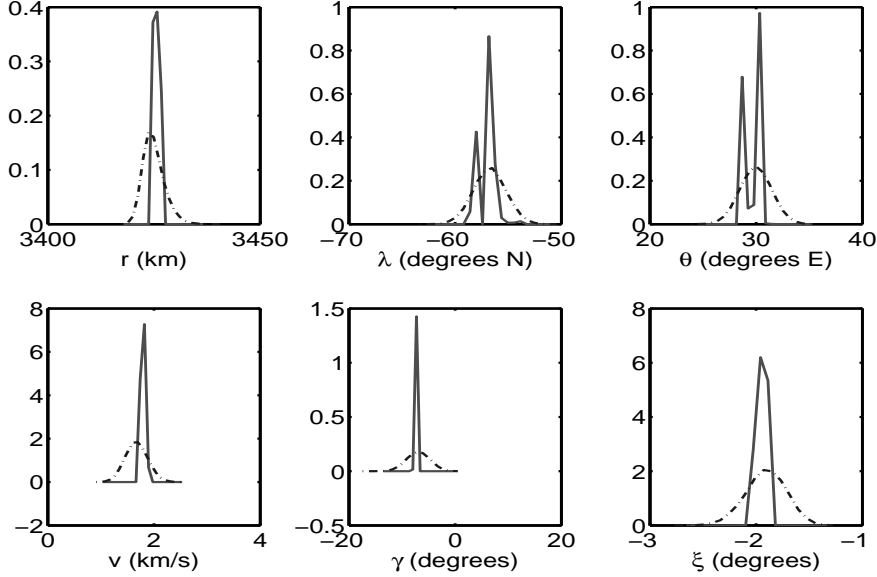
Fig. 9. Final posterior unidimensional marginal PDF for all states obtained from KLPF estimator (solid line) with 3000 particles and particle filter (hyphenated line) with 50000 particles.

the KLPF uncertainty propagation methodology, which is then applied to estimate states of a nonlinear system and its performance is compared with particle filtering based estimation techniques.

It is observed that the developed estimation technique is able to predict the evolution of posterior densities better than particle filters. Moreover, for the hypersonic reentry problem, the KLPF based estimator outperformed the sequential Monte-Carlo based methods in terms of reducing the variance and capturing the localization of uncertainty. We conclude that explicitly accounting prior dynamics via KLPF formulation, enables substantial filtering performance improvement over commonly used particle filtering technique.

REFERENCES

[1] H. Risken, *The Fokker-Planck equation: Methods of solution and applications*. Springer Verlag, 1996, vol. 18.

[2] M. Wehner and W. Wolfer, "Numerical evaluation of path-integral solutions to Fokker-Planck equations," *Physical Review A*, vol. 27, no. 5, pp. 2663–2670, 1983.

[3] G. Wei, "A unified approach for the solution of the Fokker-Planck equation," *Journal of Physics A: Mathematical and General*, vol. 33, pp. 4935–4953, 2000.
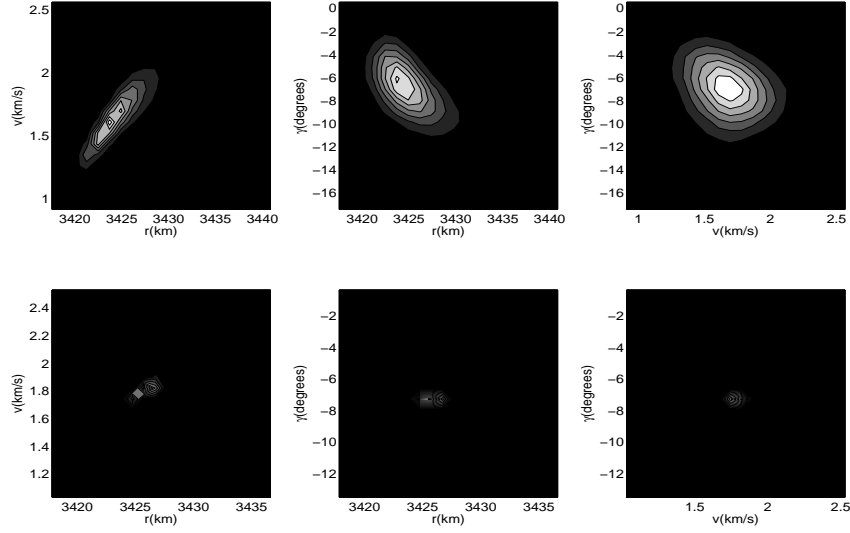
(a) Top row- particle filter, bottom row- KLPF estimator.
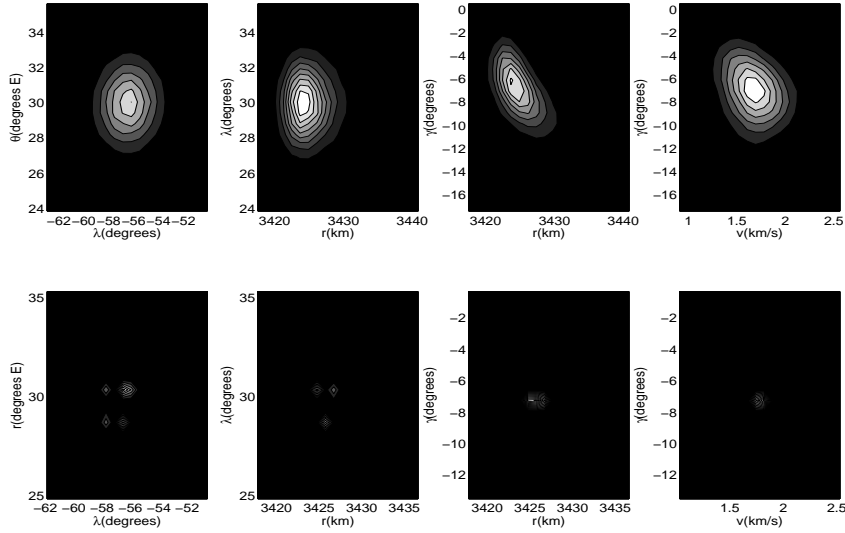
(b) Top row- particle filter, bottom row- KLPF estimator.

Fig. 10. Plots for the final posterior bivariate marginal PDFs obtained from KLPF estimator with 3000 particles and particle filter with 50000 particles. The darker regions represent lower PDF value and the lighter regions higher.

[4] M. Kumar, P. Singla, S. Chakravorty, and J. Junkins, "The partition of unity finite element approach to the stationary Fokker-Planck equation," in *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*. Keystone, CO: AIAA, August 2006, paper 2006-6285.

[5] J. Rust, "Using randomization to break the curse of dimensionality," *Econometrica*, vol. 65, no. 3, pp. 487–516, 1997.

[6] F. Daum, J. Huang, and R. Co, "Curse of dimensionality and particle filters," in *Proceedings of IEEE Aerospace Conference*, 2003, pp. 4_1979–4_1993.

[7] A. Halder and R. Bhattacharya, "Beyond Monte Carlo: A computational framework for uncertainty propagation in planetary entry, descent and landing," in *AIAA Guidance, Navigation and Control Conference, Toronto, ON*, 2010.

[8] K. Kikuchi, M. Yoshida, T. Maekawa, and H. Watanabe, "Metropolis Monte Carlo method as a numerical technique to solve the Fokker–Planck equation," *Chemical Physics Letters*, vol. 185, no. 3-4, pp. 335–338, 1991.

[9] M. Kumar, S. Chakravorty, and J. Junkins, "A semianalytic meshless approach to the transient Fokker-Planck equation," *Probabilistic Engineering Mechanics*, vol. 25, no. 3, pp. 323–331, 2010.

[10] J. Garcke and M. Griebel, "On the computation of the eigenproblems of hydrogen and helium in strong magnetic and electric fields with the sparse grid combination technique," *Journal of Computational Physics*, vol. 165, no. 2, pp. 694–716, 2000.

[11] R. Bellman, *Dynamic Programming*. Princeton, NJ: Princeton Univ. Press, 1957.

[12] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *Signal Processing, IEEE Transactions on*, vol. 50, no. 2, pp. 174–188, 2002.

[13] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[14] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," in *IEE Proceedings on Radar and Signal Processing*, vol. 140, no. 2, 1993, pp. 107–113.

[15] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman filter: Particle filters for tracking applications*. Artech House Publishers, 2004.

[16] J. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *Journal of the American Statistical Association*, vol. 93, no. 443, pp. 1032–1044, 1998.

[17] C. Snyder, T. Bengtsson, P. Bickel, and J. Anderson, "Obstacles to high-dimensional particle filtering," *Monthly Weather Review*, vol. 136, no. 12, pp. 4629–4640, 2008.

[18] G. Casella and C. Robert, "Rao-Blackwellisation of sampling schemes," *Biometrika*, vol. 83, no. 1, pp. 81–94, 1996.

[19] K. Karhunen, "Über lineare methoden in der wahrscheinlichkeitsrechnung," *Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys*, vol. 37, pp. 1–79, 1947.

[20] A. Lasota and M. Mackey, *Chaos, fractals, and noise: Stochastic aspects of dynamics*. Springer, 1994, vol. 97.

[21] A. Halder and R. Bhattacharya, "Dispersion analysis in hypersonic flight during planetary entry using stochastic Liouville equation," *Journal of Guidance, Control and Dynamics*, vol. 34, no. 2, pp. 459–474, 2011.

[22] L. Evans, "Partial differential equations," *Graduate Studies in Mathematics*, vol. 19, 1998.

[23] P. Dutta and R. Bhattacharya, "Hypersonic state estimation using the Frobenius-Perron operator," *Journal of Guidance, Control and Dynamics*, vol. 34, no. 2, pp. 325–344, 2011.

[24] F. Daum and M. Krichman, "Non-particle filters," in *Proceedings of SPIE*, vol. 6236, 2006, p. 623614.

[25] P. Dutta, A. Halder, and R. Bhattacharya, "Uncertainty quantification for stochastic nonlinear systems using Perron-

Frobenius operator and Karhunen-Loève expansion," in *IEEE Multiconference in Systems and Control*.   IEEE, 2012, pp. 1449–1454.

[26] ——, "Nonlinear filtering with transfer operator," in *American Control Conference (ACC), 2013*.   IEEE, 2013, pp. 3069–3074.

[27] M. Loéve, "Fonctions aleatories du second ordre. Supplement to P. Lévy, Processus Stochastic et Mouvement Brownien, Gauthier-Villars, Paris," 1948.

[28] M. Grigoriu, *Stochastic calculus: Applications in science and engineering*.   Birkhauser, 2002.

[29] A. Leon-Garcia, *Probability, Statistics, and Random Processes for Electrical Engineering*.   Prentice Hall, 1994.

[30] R. Ghanem and P. Spanos, *Stochastic finite elements: A spectral approach*.   Dover Pubns, 2003.

[31] B. Øksendal, *Stochastic differential equations: An introduction with applications*.   Springer Verlag, 2003.

[32] P. Frauenfelder, C. Schwab, and R. Todor, "Finite elements for elliptic problems with stochastic coefficients," *Computer methods in applied mechanics and engineering*, vol. 194, no. 2-5, pp. 205–228, 2005.

[33] M. Abramowitz and I. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*.   Dover publications, 1964, vol. 55, no. 1972.

[34] C. Schwab and R. Todor, "Karhunen-Loève approximation of random fields by generalized fast multipole methods," *Journal of Computational Physics*, vol. 217, no. 1, pp. 100–122, 2006.

[35] A. Longtin, "Stochastic delay-differential equations," *Complex Time-Delay Systems*, pp. 177–195, 2010.

[36] T. Frank, "Delay Fokker-Planck equations, perturbation theory, and data analysis for nonlinear stochastic systems with time delays," *Physical Review E*, vol. 71, no. 3, p. 031106, 2005.

[37] C. Baker, C. Paul, and D. Willé, "Issues in the numerical solution of evolutionary delay differential equations," *Advances in Computational Mathematics*, vol. 3, no. 1, pp. 171–196, 1995.

[38] A. Bryson and Y. Ho, *Applied optimal control: Optimization, estimation, and control*.   Hemisphere Pub, 1975.

[39] D. J. Higham, X. Mao, and A. M. Stuart, "Strong convergence of euler-type methods for nonlinear stochastic differential equations," *SIAM Journal on Numerical Analysis*, vol. 40, no. 3, pp. 1041–1063, 2002.

[40] L. Rüschendorf, "The Wasserstein distance and approximation theorems," *Probability Theory and Related Fields*, vol. 70, no. 1, pp. 117–129, 1985.

[41] J. Cuesta and C. Matran, "Notes on the Wasserstein metric in Hilbert spaces," *The Annals of Probability*, pp. 1264–1276, 1989.

[42] C. Villani, *Topics in optimal transportation*.   AMS Bookstore, 2003, vol. 58.

[43] A. Halder and R. Bhattacharya, "Model validation: A probabilistic formulation," in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, December 2011, pp. 1692 –1697.

[44] V. Beneš, "Exact finite-dimensional filters for certain diffusions with nonlinear drift," *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 5, no. 1-2, pp. 65–92, 1981.

[45] A. Bain and D. Crişan, *Fundamentals of stochastic filtering*.   Springer Verlag, 2008, vol. 60.

[46] D. Crisan, "A direct computation of the Beneš filter conditional density," *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 55, no. 1-2, pp. 47–54, 1995.

[47] S. Chib and E. Greenberg, "Understanding the Metropolis-Hastings algorithm," *American Statistician*, pp. 327–335, 1995.

[48] J. Chern and N. Vinh, "Optimum reentry trajectories of a lifting vehicle," NASA, Tech. Rep. NASA-CR-3236, 1980.

[49] K. Bollino, I. Ross, and D. Doman, "Optimal nonlinear feedback guidance for reentry vehicles," in *AIAA Guidance, Navigation and Control Conference*, 2006.